Home Work 1

Due Date: 12th September 2019 (8am, in class).
(Attempt all questions)

Q. 1]
(i) Consider the following subset of $\mathbb{R}^2$: $\mathcal{Q} = \{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 \ : \ 9x_1^2 + 16x_1x_2 + 21x_2^2 \leq 16\}$. Say whether $\mathcal{Q}$ is closed and bounded. If $\mathcal{Q}$ is bounded then find $M > 0$ & $\hat{\mathbf{x}} \in \mathcal{Q}$ such that $\|\mathbf{x}\|_2 \leq M$ for all $\mathbf{x} \in \mathcal{Q}$ and $\|\hat{\mathbf{x}}\|_2 = M$. If $\mathcal{Q}$ is not bounded then find $\hat{\mathbf{x}} \in \mathcal{Q}$ such that $\|\hat{\mathbf{x}}\|_2 > N$, where N is an arbitrary big positive number. . [5 marks]

(ii) Consider the following subset in the space of 2×2 real matrices: $\mathcal{S} = \{X \in \mathbb{R}^{2\times 2} :$ X is symmetric & positive semidefinite, $\text{trace}(X) = 1\}$. Say whether $\mathcal{S}$ is closed and bounded. If $\mathcal{S}$ is bounded then find $M > 0$ & $\hat{X} \in \mathcal{S}$ such that $\|X\|_F \leq M$ for all $X \in \mathcal{S}$ and $\|\hat{X}\|_F = M$. Here $\|\cdot\|_F$ denotes the Frobenius norm of a matrix. [5 marks]

Ans: (Present your answer in tabular form given below)

	Closed?	Bounded ?	M	$\hat{\mathbf{x}}$ or $\hat{X}$
Q. 1(i)	YES	YES	$\frac{4}{\sqrt{5}}$	$(-\frac{8}{5}, \frac{4}{5})$ or $(\frac{8}{5}, -\frac{4}{5})$
Q. 1(ii)	YES	YES	1	$\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ $a \geq 0,\ c \geq 0,\ a + c = 1,$ $b^2 \leq ac,\ a^2 + 2b^2 + c^2 = 1.$

Q. 2] Let β denote the number consisting of last three digits of your S.R. number and set $\epsilon = (0.001\beta + 1)10^{-4}$. Consider an iterative algorithm for solving the optimization problem $f^* = \min_{\mathbf{x}} f(\mathbf{x})$. Let $\mathbf{x}_k$ be the solution produced by the algorithm after k-th iteration and $e_k = f(\mathbf{x}_k) - f^*$ is the corresponding error. For the following cases plot e_k vs k and calculate the minimum number of iterations k required to reduce e_k below ϵ, 0.1ϵ, 0.01ϵ and 0.0001ϵ:

- [Sublinear Convergence:] (i) $e_k = 0.5/\sqrt{k}$, (ii) $e_k = 0.5/k$, (iii) $e_k = 0.5/k^2$,
- [Linear Convergence:] (iv) $e_1 = 0.5$, $e_k = 0.975e_{k-1}, k \geq 2$,
- [Quadratic Convergence:] (v) $e_1 = 0.5$, $e_k = 1.9e_{k-1}^2, k \geq 2$.

. [10 marks]

Ans: Attach the following plots:

- e_k (log scale) vs k (log scale) for (i), (ii) & (iii) together in one plot.
- e_k (log scale) vs k (normal scale) for (iv).
- e_k (log scale) vs k (normal scale) for (v).

$\beta =$

$\epsilon \ =$

Fill the blanks in the below table with N_δ (for $\delta = \epsilon,\ 0.1\epsilon, \ldots$) which is the minimum number of iterations needed such that error $e_k \leq \delta,\ \forall k \geq N_\delta$.

$\delta =$	$\delta = \epsilon$	$\delta = 0.1\epsilon$	$\delta = 0.01\epsilon$	$\delta = 0.0001\epsilon$
Q. 2(i)	$\left\lceil \frac{0.25}{\delta^2} \right\rceil$			
Q. 2(ii)	$\left\lceil \frac{0.5}{\delta} \right\rceil$			
Q. 2(iii)	$\left\lceil \sqrt{\frac{0.5}{\delta}} \right\rceil$			
Q. 2(iv)	$\left\lceil \frac{\log(2\,\delta)}{\log(0.975)} + 1 \right\rceil$			
Q. 2(v)	$\left\lceil \log_2\left(\frac{\log(1.9\,\delta)}{\log(0.95)}\right) + 1 \right\rceil$			

Q. 3(i)
Let $f : \mathbb{R}^4 \to \mathbb{R}$ be twice continuously differentiable function. The largest absolute eigen value of the Hessian matrix of f at all points is bounded above by 25. We are given that $\mathbf{x}_0 = [4,\ 0,\ -2,\ 1]^\top$, $f(\mathbf{x}_0) = 6$ and $\nabla f(\mathbf{x}_0) = [8,\ 4,\ 4,\ 2]^\top$. Using 2nd order Taylor Series find a quadratic function $g : \mathbb{R}^4 \to \mathbb{R}$ such that $g(\mathbf{x}_0) = f(\mathbf{x}_0)$ and $f(\mathbf{x}) \leq g(\mathbf{x})\, \forall \mathbf{x} \in \mathbb{R}^4$. Can the minimum value of f be more than 4? If the answer is no then find $\tilde{\mathbf{x}}$ such that $f(\tilde{\mathbf{x}}) \leq 4$. If the answer is yes then find $\tilde{\mathbf{x}}$ such that $4 < f(\tilde{\mathbf{x}}) < 6$. [4+3 marks]
Hint: As g is quadratic, it can be expressed as: $g(\mathbf{x}) = \mathbf{x}^\top A\mathbf{x} + \mathbf{b}^\top \mathbf{x} + c$, where $A \in \mathbb{R}^{4\times 4}$, $\mathbf{b} \in \mathbb{R}^4$ and $c \in \mathbb{R}$. Therefore, you need to simply specify the values of $A, \mathbf{b}, c$ for your g.

Ans: $A = \frac{25}{2}\mathbf{I}, \quad b = (-92,\ 4,\ 54,\ -23), \quad c = 242.5$
No, minimum value of f can not be more than 4.

$$\tilde{\mathbf{x}} \ = \ -\frac{1}{25}(-92,\ 4,\ 54,\ -23) \ = \ (3.68,\ -0.16,\ -2.16,\ 0.92).$$

and $\quad f(\tilde{\mathbf{x}}) \leq g(\tilde{\mathbf{x}}) = 4.$

Q. 3(ii)

Consider the function $f(\mathbf{x}) = (\mathbf{p}^\top\mathbf{x})^2 + (\mathbf{q}^\top\mathbf{x})^2$, $\mathbf{p}, \mathbf{q}, \mathbf{x} \in \mathbb{R}^3$. Assume $\mathbf{p} \neq 0$, $\mathbf{q} \neq 0$, $\mathbf{p} + \mathbf{q} \neq 0$. Two friends Siri and Cortona have opinions about f. Siri claims that there is an $\mathbf{x} \in \mathbb{R}^3$ with $\|\mathbf{x}\|_2 = 1$ but $f(\mathbf{x}) = 0$. Cortona says f is coercive. What can you say about f: (a) Both claims are correct, (b) Siri is correct and Cortona's claim is false, (c) Siri's claim is false and Cortona is correct, (d) Both claims are false.
. [3 marks]

Ans. (b) Siri is correct and Cortona's claim is false.

Q. 4]

Recall that $\bar{\mathbf{x}}$ is a local maxima of f if the function value does not increase (decrease for local minima) when you move very small distance from $\bar{\mathbf{x}}$ in any direction. We call $\bar{\mathbf{x}}$ a saddle point of f if the function value increases when we move from $\bar{\mathbf{x}}$ in some direction and descreases when we move from $\bar{\mathbf{x}}$ in another direction. Let $\mathbf{d}_\theta = [\cos(\theta), \sin(\theta)]^\top$ donote the direction vector with angle $\theta \in [0, 2\pi]$.

(i) Consider the function $f(\mathbf{x}) = 10x_1^2+10x_1x_2+x_2^2+4x_1-10x_2+2$ over $[-5,5]\times[-5,5]$. Plot the function and see whether it has a local maxima / minima / saddle point at $\bar{\mathbf{x}} = (1.8, -4)$. For clarity, plot $[f(\bar{\mathbf{x}}+\alpha\mathbf{d}_\theta) - f(\bar{\mathbf{x}})]$ vs θ for $\alpha = 0.01$ and see whether it is always non-negative (if local minima), always non-positive (if local maxima) or has both positive and negative (if saddle point). Also, calculate $\nabla f(\bar{\mathbf{x}})$ and eigen values of $\nabla^2 f(\bar{\mathbf{x}})$. [5 marks]

(ii) Repeat part (i) with $f(\mathbf{x}) = 16x_1^2+8x_1x_2+10x_2^2+12x_1-6x_2+2$, $\bar{\mathbf{x}} = (-0.5, 0.5)$.
. [5 marks]

Ans: Attach the plots and fill up the following table:

	$\nabla f(\bar{\mathbf{x}})$	eigen values of $\nabla^2 f(\bar{\mathbf{x}})$	at $\bar{\mathbf{x}}$ local maxima/minima/saddle point?
Q. 4(i)	(0, 0)	$24.4536, \ -2.4536$	saddle point
Q. 4(ii)	(0, 0)	$36, \ 16$	local minima

Q. 5] Consider the following 1-dimensional minimization problem:

$$\min_{x \in [a,b]} f(x)\,. \tag{1}$$

We are given a, b, a subroutine named "zoof" (Zero Order Oracle for f). The subroutine "zoof" is suppose to return only the value of f for any given $x \in [a, b]$ i.e. "fvalueAtx = zoof(x)". Let x^* be the unique minimizer of the problem (1). Given a tolerance value $\varepsilon > 0$, we are interested in finding $\hat{x} \in [a, b]$ such that $|\hat{x} - x^*| \leq \varepsilon$. Note that x^* and the function f are unknown but we are allowed to call the "zoof" subroutine multiple times. Here is a generic pseudo code to find $\hat{x}$:

- Initialize: $x_l = a$, $x_u = b$.
- In loop:
  - $d = (x_u - x_l) * \rho$. [$0 < \rho < 1$ is an algorithm specific choice]
  - $x_- = x_u - d,\ x_+ = x_l + d$.
  - If zoof(x_-) $<$ zoof(x_+) then $x_u = x_+$ otherwise $x_l = x_-$.
- Output: $\hat{x} = 0.5(x_l + x_u)$, tolerance $= 0.5 * (x_u - x_l)$, NSC= number of times "zoof" subroutine was called.

We will now specialize the above generic structure to describe two line search techniques. In Golden search (GS) we set ρ to a fixed value $\lambda/(1 + \lambda)$, where $\lambda = 0.5 * (1 + \sqrt{5})$. Now, implement Golden Search as a function (in Matlab / Python) with structure:

$$[\hat{x},\ \text{tolerance},\ \text{NSC}\,] \;=\; \text{GS}(a, b,\ \text{zoof},\ \varepsilon).$$

Note that the "loop" part of the GS algorithm should continue until tolerance $\leq \varepsilon$ is satisfied.

Similarly, implement Fibonacci search (FS) as a function:

$$[\hat{x},\ \text{tolerance},\ \text{NSC}\,] \;=\; \text{FS}(a, b,\ \text{zoof},\ N).$$

Here the "loop" part should be executed only $(N - 1)$ times and at k-th iteration use $\rho = \frac{F_{N-k}}{F_{N-k+1}}$, where $F_0 = 1$, $F_1 = 1$, $F_n = F_{n-1} + F_{n-2}, n \geq 2$ is the Fibonacci sequence.

Now, implement two different "zoof" subroutine such that
(i) zoof(x) returns $|2 - x| + |5 - 4x| + |8 - 9x|$,
(ii) zoof(x) returns $3(x - 1)^2 - e^{x-1}$.

Fix $a = 0$, $b = 3$. Now, for each of the above choices of "zoof" first call the FS subroutine with $N = 5$, $N = 10$, and $N = 20$. Then for each case call the GS subroutine with ε set to the tolerance value returned by the FS subroutine. Report $\hat{x}$, tolerance, NSC for all the runs. Also, plot the functions (i) and (ii) over the interval $[a, b]$. Note that the above line search subroutines do not need differentiability of f: (ii) is differentiable where as (i) is not. The crucial property that the above line search methods need is that f has a unique local minima over $[a, b]$.

Reference: Details on the above line search techniques can be found in section 2.1 of this book.

. [5 + 5 marks]

Ans: You do not need to submit the code. But keep the code with you as we will use those functions in later assignments as well. Present the output of your line search subroutines as follows:

case	FS subroutine			GS subroutine		
	$\hat{x}$	tolerance	NSC	$\hat{x}$	tolerance	NSC
(i)zoof, $N = 5$	1.3125	0.1875	8	0.8435	0.1353	10
(i)zoof, $N = 10$	0.9270	0.0169	18	0.8876	0.0122	20
(i)zoof, $N = 20$	0.8892	0.0001	38	0.8889	0.0001	40
(ii)zoof, $N = 5$	1.3125	0.1875	8	1.2812	0.1353	10
(ii)zoof, $N = 10$	1.2303	0.0169	18	1.1976	0.0122	20
(ii)zoof, $N = 20$	1.2047	0.0001	38	1.2045	0.0001	40